

The listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1.-22. (Cancelled)

23. (Currently Amended) A speech synthesis device, the device comprising:

a first storage means for storing a plurality of pieces of voice unit data representative of one or more speech words;

a selection means for selecting voice unit data whose reading is common with a speech word composing inputted sentence information from the plurality of pieces of voice unit data stored in the first storage means;

a missing part synthesis means, for a speech word among the sentence information for which the selection means could not select the voice unit data, for synthesizing speech data representative of a desired speech waveform; and

a synthesis means for combining the voice unit data selected from the selection means and the speech data synthesized by the missing part synthesis means to create data representative of a synthesis speech corresponding to the sentence information,

wherein the missing part synthesis means has a second storage means for storing a plurality of pieces of data representative of one or more pitches of voice waveform fragments, the one or more pitches of voice waveform fragments being cut off in a unit of voice pitch from an actual speech waveform; and

wherein data representative of voice waveform fragments composing the speech word whose voice unit data could not be selected is acquired from the second storage means and the acquired data is mutually combined to synthesize the speech data representative of the desired speech waveform.

24. (Previously Presented) The speech synthesis device according to claim 23, further comprising a cadence prediction means for predicting a cadence of the speech word composing the inputted sentence information, wherein the selection means selects voice unit data whose cadence matches with a cadence prediction result under predetermined conditions.

25. (Previously Presented) The speech synthesis device according to claim 24, wherein the selection means operates to exclude from the objects of selection voice unit data whose cadence does not match with the cadence prediction result under the predetermined conditions.

26. (Previously Presented) The speech synthesis device according to claim 24, wherein the missing part synthesis means comprises a missing part cadence prediction means that predicts the cadence of the speech word for which the selection means could not select voice unit data, and

wherein the synthesis means identifies a phoneme and acquires data representative of the voice unit data composing the speech word, for which the selection means could not select voice unit data and acquires from the second storage means, converts the acquired data such that the phoneme or the speech waveform fragment represented by the data matches with the cadence result predicted by the missing part cadence prediction means, and combines the converted data to synthesize speech data representative of the desired speech waveform.

27. (Previously Presented) The speech synthesis device according to claim 25, wherein the missing part synthesis means comprises a missing part cadence prediction means that predicts the cadence of the speech word for which the selection means could not select voice unit data, and

wherein the synthesis means identifies a phoneme and acquires data representative of the voice unit data composing the speech word, for which the selection means could not select voice unit data and acquires from the second storage means, converts the acquired data such that the phoneme or the speech waveform fragment represented by the data matches with the cadence result predicted by the missing part cadence prediction means, and combines the converted data to synthesize speech data representative of the desired speech waveform.

28. (Previously Presented) The speech synthesis device according to claim 24, wherein the first storage means stores cadence data representative of time variations in a pitch of a voice unit represented by voice unit data with the cadence data being associated with the voice unit data, and

wherein the selection means selects, from the respective voice unit data, voice unit data whose reading is common with the speech word composing the sentence information and for which a time variation in the pitch represented by the associated cadence data is closest to the cadence prediction result.

29. (Previously Presented) The speech synthesis device according to claim 25, wherein the first storage means stores cadence data representative of time variations in a pitch of a voice unit represented by voice unit data with the cadence data being associated with the voice unit data, and

wherein the selection means selects, from the respective voice unit data, voice unit data whose reading is common with the speech word composing the sentence information and for which a time variation in the pitch represented by the associated cadence data is closest to the cadence prediction result.

30. (Previously Presented) The speech synthesis device according to claim 23, wherein the device further comprises utterance speed conversion means for acquiring

utterance speed data specifying conditions of a speed for uttering the synthetic speech and selects or converts speech data and/or voice unit data composing data representative of the synthetic speech such that the speech data and/or voice unit data represents speech that is uttered at a speed fulfilling the conditions specified by the utterance speed data.

31. (Previously Presented) The speech synthesis device according to claim 24, wherein the device further comprises utterance speed conversion means for acquiring utterance speed data specifying conditions of a speed for uttering the synthetic speech and selects or converts speech data and/or voice unit data composing data representative of the synthetic speech such that the speech data and/or voice unit data represents speech that is uttered at a speed fulfilling the conditions specified by the utterance speed data.

32. (Previously Presented) The speech synthesis device according to claim 25, wherein the device further comprises utterance speed conversion means for acquiring utterance speed data specifying conditions of a speed for uttering the synthetic speech and selects or converts speech data and/or voice unit data composing data representative of the synthetic speech such that the speech data and/or voice unit data represents speech that is uttered at a speed fulfilling the conditions specified by the utterance speed data.

33. (Previously Presented) The speech synthesis device according to claim 30, wherein the utterance speed conversion means, by eliminating a segment representing a speech waveform fragment from speech data and/or voice unit data composing data representative of the synthetic speech or adding a segment representative of a speech waveform fragment to the voice unit data and/or speech data, converts the voice unit data and/or speech data such that the voice unit data and/or speech data represents

speech that is uttered at a speed fulfilling the conditions specified by the utterance speed data.

34. (Currently Amended) A speech synthesis device, the device comprising:

a first storage means for storing a plurality of pieces of voice unit data representative of one or more speech words;

a selection means for selecting voice unit data whose reading is common with a speech word composing inputted sentence information from the plurality of pieces of voice unit data stored in the first storage means;

a missing part synthesis means, for a speech word among the sentence information for which the selection means could not select the voice unit data, for synthesizing speech data representative of a desired speech waveform by mutually combining one or more pitches of voice waveform fragments cut off in a unit of voice pitch from an actual speech waveform; and

a synthesis means for combining the voice unit data selected from the selection means and the speech data synthesized by the missing part synthesis means to create data representative of a synthesis speech corresponding to the sentence information,

wherein the first storage means stores phonetic data representative of a reading of the voice unit data with the phonetic data being associated with the voice unit data, and

wherein the selection means operates to handle voice unit data which is associated with phonetic data representative of a reading matching with the reading of the speech word composing the sentence information as voice unit data whose reading is common with the speech word.

35. (Currently Amended) A speech synthesis method, the method comprising the steps of:

storing a plurality of pieces of voice unit data representative of one or more speech words in a first memory;

selecting voice unit data whose reading is common with a speech word composing inputted sentence information from the plurality of pieces of voice unit data stored in the first memory;

synthesizing a missing part, for a speech word among the sentence information for which the voice unit data could not be selected in the selecting step, by synthesizing speech data representative of a desired speech waveform; and

combining the voice unit data selected from the selection means and the speech data synthesized in the missing part synthesizing step to create data representative of a synthesis speech corresponding to the sentence information,

wherein the missing part synthesizing step stores a plurality of pieces of data representative of one or more pitches of voice waveform fragments using a second memory, the one or more pitches of voice waveform fragments being cut off in a unit of voice pitch from an actual speech waveform; and

wherein data representative of voice waveform fragments composing the speech word whose voice unit data could not be selected is acquired from the second memory and the acquired data is combined to synthesize the speech data representative of the desired speech waveform.

36. (Currently Amended) A speech synthesis method, the method comprising the steps of:

storing a plurality of pieces of voice unit data representative of one or more speech words in a first memory;

selecting voice unit data whose reading is common with a speech word composing inputted sentence information from the plurality of pieces of voice unit data stored in the first memory;

synthesizing a missing part, for a speech word among the sentence information for which the selection means could not select the voice [[units]] unit data, by synthesizing speech data representative of a desired speech waveform by mutually combining one or more pitches of voice waveform fragments cut off in a unit of voice pitch from an actual speech waveform; and

combining the voice unit data selected from the selection means and the speech data synthesized in the missing part synthesis step to create data representative of a synthesis speech corresponding to the sentence information,

wherein the first memory stores phonetic data representative of a reading of the voice unit data with the phonetic data being associated with the voice unit data, and

wherein the selecting step handles voice unit data which is associated with phonetic data representative of a reading matching with the reading of the speech word composing the sentence information as voice unit data whose reading is common with the speech word.

37. (Currently Amended) A computer readable medium recording a computer program, the computer program causing a computer to operate as:

a first storage means for storing a plurality of pieces of voice unit data representative of one or more speech words;

a selection means for selecting voice unit data whose reading is common with a speech word composing inputted sentence information from the plurality of pieces of voice unit data stored in the first storage means;

a missing part synthesis means, for a speech word among the sentence information for which the selection means could not select the voice [[units]] unit data, for synthesizing speech data representative of a desired speech waveform; and

a synthesis means for combining the voice unit data selected from the selection means and the speech data synthesized by the missing part synthesis means to create data representative of a synthesis speech corresponding to the sentence information,

wherein the missing part synthesis means has a second storage means for storing a plurality of pieces of data representative of one or more pitches of voice waveform fragments, the one or more pitches of voice waveform fragments being cut off in a unit of voice pitch from an actual speech waveform; and

wherein data representative of voice waveform fragments composing the speech word whose voice unit data could not be selected is acquired from the second storage means and the acquired data is mutually combined to synthesize the speech data representative of the desired speech waveform.

38. (Currently Amended) A computer readable medium recording a computer program, the computer program causing a computer to operate as:

a first storage means for storing a plurality of pieces of voice unit data representative of one or more speech words;

a selection means for selecting voice unit data whose reading is common with a speech word composing inputted sentence information from the plurality of pieces of voice unit data stored in the first storage means;

a missing part synthesis means, for a speech word among the sentence information for which the selection means could not select the voice [[units]] unit data, for synthesizing speech data representative of a desired speech waveform by mutually combining one or more pitches of voice waveform fragments cut off in a unit of voice pitch from an actual speech waveform; and

a synthesis means for combining the voice unit data selected from the selection means and the speech data synthesized by the missing part synthesis means to create data representative of a synthesis speech corresponding to the sentence information,

wherein the first storage means stores phonetic data representative of a reading of the voice unit data with the phonetic data being associated with the voice unit data, and

wherein the selection means operates to handle voice unit data which is associated with phonetic data representative of a reading matching with the reading of the speech word composing the sentence information as voice unit data whose reading is common with the speech word.

39. (Previously Presented) The speech synthesis device according to claim 24, wherein the missing part synthesis means comprises a missing part cadence prediction means that predicts the cadence of the speech word for which the selection means could not select voice unit data, and

wherein the synthesis means identifies a phoneme and acquires data representative of the voice unit data composing the speech word, for which the selection means could not select voice unit data and acquires from the second storage means, converts the acquired data such that the phoneme or the speech waveform fragment represented by the data matches with the cadence result predicted by the missing part cadence prediction means, and combines the converted data to synthesize speech data representative of the desired speech waveform.

40. (Previously Presented) The speech synthesis device according to claim 23, wherein the first storage means stores cadence data representative of time variations in a pitch of a voice unit represented by voice unit data with the cadence data being associated with the voice unit data, and

wherein the selection means selects, from the respective voice unit data, voice unit data whose reading is common with the speech word composing the sentence information and for which a time variation in the pitch represented by the associated cadence data is closest to the cadence prediction result.

41. (Canceled)